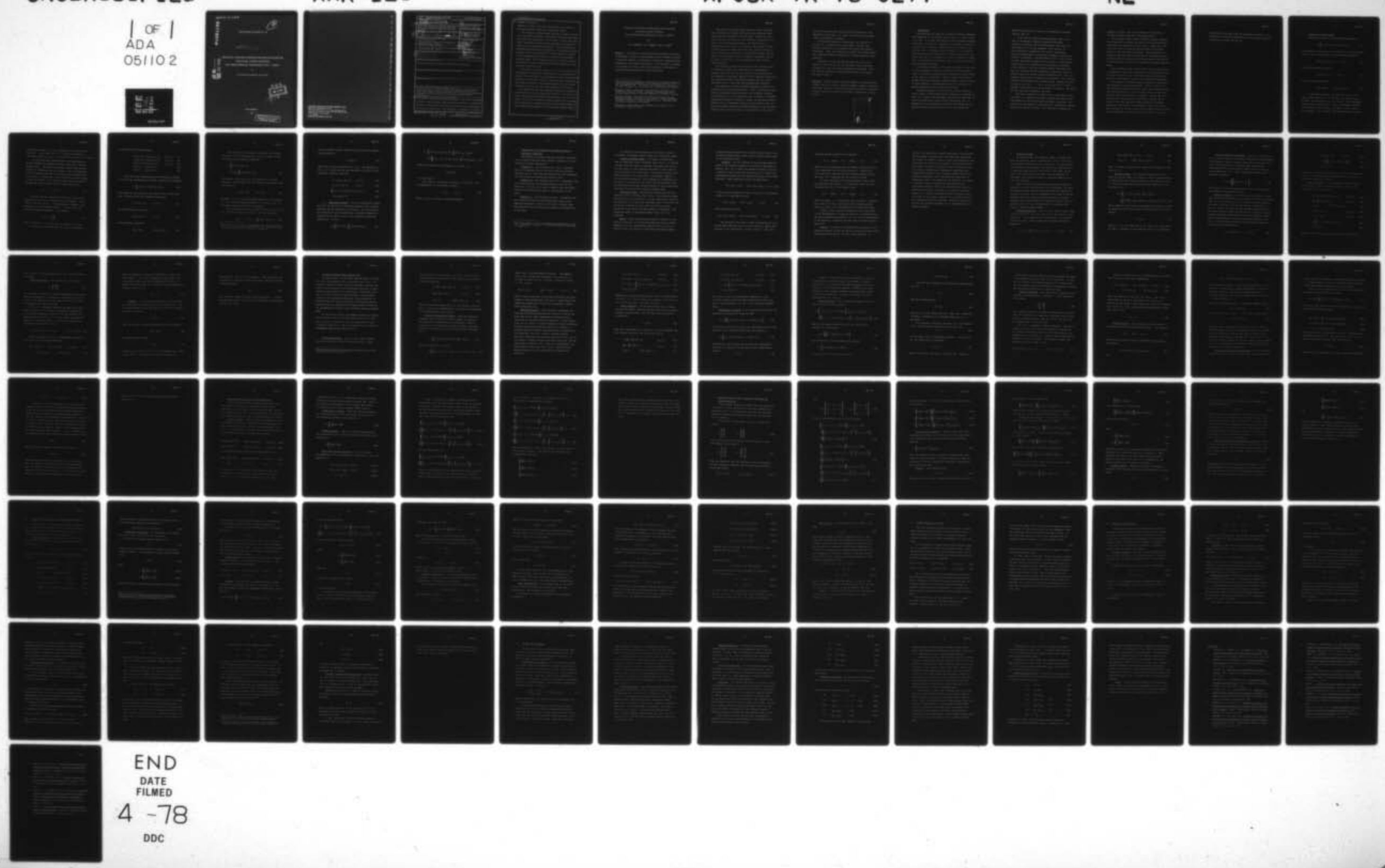
AD-A051 102 RICE UNIV HOUSTON TEX AERO-ASTRONAUTICS GROUP F/G 12/1
SEQUENTIAL CONJUGATE GRADIENT-RESTORATION ALGORITHM FOR OPTIMAL--ETC(U)
1977 J R CLOUTIER, B P MOHANTY, A MIELE AFOSR-76-3075
UNCLASSIFIED AAR-126 AFOSR-TR-78-0277 NL

1 OF 1
ADA
051102



END
DATE
FILMED
4 -78
DDC

2

AERO-ASTRONAUTICS REPORT NO. 126

SEQUENTIAL CONJUGATE GRADIENT-RESTORATION ALGORITHM FOR OPTIMAL CONTROL PROBLEMS WITH NONDIFFERENTIAL CONSTRAINTS, PART 1, THEORY

by

J.R. CLOUTIER, B.P. MOHANTY, and A. MIELE

RICE UNIVERSITY

1977

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFOSR-TR- 78 - 0277 (18) (19) / 2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER (9)
4. TITLE (and Subtitle) (6) SEQUENTIAL CONJUGATE GRADIENT-RESTORATION ALGORITHM FOR OPTIMAL CONTROL PROBLEMS WITH NONDIFFERENTIAL CONSTRAINTS, PART 1, THEORY.	5. TYPE OF REPORT & PERIOD COVERED Interim rept., / 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) J.R. Cloutier, B.P. Mohanty [redacted] A. Miele	8. CONTRACT OR GRANT NUMBER(s) (15) AFOSR-76-3075 new
9. PERFORMING ORGANIZATION NAME AND ADDRESS Rice University Department of Mechanical Engineering Houston, Texas 77001	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (16) 61102F 2304/A3 (17)
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, DC 20332	12. REPORT DATE (11) 1977 / 13. NUMBER OF PAGES ~~79~~ (12) 23p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (14) AAR-126	15. SECURITY CLASS. (of this report) UNCLASSIFIED / 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Optimal control, gradient methods, conjugate-gradient methods, numerical methods, computing methods, gradient-restoration algorithms, sequential gradient-restoration algorithms, sequential conjugate gradient-restoration algorithm, nondifferential constraints.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A sequential conjugate gradient-restoration algorithm is developed in order to solve optimal control problems involving a functional subject to differential constraints, nondifferential constraints, and terminal constraints. The algorithm is composed of a (OVER)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

402 169

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (Continued)

sequence of cycles, each cycle consisting of two phases, a conjugate gradient phase and a restoration phase.

The conjugate gradient phase involves a single iteration and is designed to decrease the value of the functional while satisfying the constraints to first order. During this iteration, the first variation of the functional is minimized, subject to the linearized constraints. The minimization is performed over the class of variations of the control and the parameter which are equidistant from some constant multiple of the corresponding variations of the previous conjugate gradient phase. For the special case of a quadratic functional subject to linear constraints, various orthogonality and conjugacy conditions hold.

The restoration phase involves one or more iterations and is designed to restore the constraints to a predetermined accuracy, while the norm of the variations of the control and the parameter is minimized, subject to the linearized constraints. The restoration phase is terminated whenever the norm of the constraint error is less than some predetermined tolerance.

The sequential conjugate gradient-restoration algorithm is characterized by two main properties. First, at the end of each cycle, the trajectory satisfies the constraints to a given accuracy. Second, the conjugate gradient stepsize and the restoration stepsize can be chosen so that the restoration phase preserves the descent property of the conjugate gradient phase.

Sequential Conjugate Gradient-Restoration Algorithm for Optimal Control Problems with Nondifferential Constraints, Part 1, Theory[1]

by

J.R. CLOUTIER[2], B.P. MOHANTY[3], and A. MIELE[4]

Abstract. A sequential conjugate gradient-restoration algorithm is developed in order to solve optimal control problems involving a functional subject to differential constraints, nondifferential constraints, and terminal constraints. The algorithm is composed of a sequence of cycles, each cycle consisting of two phases, a conjugate gradient phase and a restoration phase.

---

[1]This research was supported by the Office of Scientific Research, Office of Aerospace Research, United States Air Force, Grant No. AF-AFOSR-76-3075. The authors are indebted to Messieurs A.K. Wu and S. Gonzalez for analytical and computational assistance.

[2]Graduate Student, Department of Mathematical Sciences, Rice University, Houston, Texas. Presently, Mathematician, Naval Surface Weapons Center, Dahlgren Laboratory, Dahlgren, Virginia.

[3]Graduate Student, Department of Mechanical Engineering and Materials Science, Rice University, Houston, Texas. Presently, Assistant Professor, School of Systems Science, Arkansas Tech University, Russelville, Arkansas.

[4]Professor of Astronautics and Mathematical Sciences, Rice University, Houston, Texas.

The conjugate gradient phase involves a single iteration and is designed to decrease the value of the functional while satisfying the constraints to first order. During this iteration, the first variation of the functional is minimized, subject to the linearized constraints. The minimization is performed over the class of variations of the control and the parameter which are equidistant from some constant multiple of the corresponding variations of the previous conjugate gradient phase. For the special case of a quadratic functional subject to linear constraints, various orthogonality and conjugacy conditions hold.

The restoration phase involves one or more iterations and is designed to restore the constraints to a predetermined accuracy, while the norm of the variations of the control and the parameter is minimized, subject to the linearized constraints. The restoration phase is terminated whenever the norm of the constraint error is less than some predetermined tolerance.

The sequential conjugate gradient-restoration algorithm is characterized by two main properties. First, at the end of each conjugate gradient-restoration cycle, the trajectory satisfies the constraints to a given accuracy; thus, a sequence of feasible suboptimal solutions is produced. Second, the conjugate gradient stepsize and the restoration stepsize can be chosen so that the restoration phase preserves the descent property of the conjugate gradient phase; thus, the value of the functional at the end of

any cycle is smaller than the value of the functional at the beginning of that cycle. Of course, restarting the algorithm might be occasionally necessary.

To facilitate numerical integrations, the interval of integration is normalized to unit length. Variable-time terminal conditions are transformed into fixed-time terminal conditions. Then, the actual time at which the terminal boundary is reached becomes a component of a vector parameter being optimized.

Convergence is attained whenever both the norm of the constraint error and the norm of the error in the optimality conditions are less than some predetermined tolerances. Several numerical examples illustrating the theory of this paper are given in Part 2.

Key Words. Optimal control, gradient methods, conjugate-gradient methods, numerical methods, computing methods, gradient-restoration algorithms, sequential gradient-restoration algorithms, sequential conjugate gradient-restoration algorithms, nondifferential constraints.

1. Introduction

Approximately ten years ago, conjugate gradient techniques began appearing on the optimal control scene. In 1967, Lasdon et al (Ref. 1) extended the conjugate gradient method developed by Fletcher and Reeves for mathematical programming problems to optimal control problems. About the same time, Horwitz and Sarachik (Ref. 2) extended Davidon's method to a real Hilbert space and applied the extension to a control problem with quadratic cost and linear constraints. Shortly thereafter, Lasdon (Ref. 3) and Tripathi and Narendra (Ref. 4) also derived extensions of Davidon's method.

One common limitation of the above algorithms is that they are not applicable directly to constrained control problems (that is, problems involving terminal constraints and/or bounds on the state or the control). However, these algorithms can handle indirectly constrained control problems, after conversion of these problems to unconstrained form; this conversion is usually achieved by means of penalty functions.

Conjugate gradient algorithms which can solve directly certain types of constrained control problems were presented in Refs. 5-8. Sinnott and Luenberger constructed an algorithm for solving problems with linear terminal constraints (Ref. 5); Heideman and Levy developed an algorithm for problems with arbitrary terminal constraints (Refs. 6-7); and Pagurek and

Woodside constructed an algorithm for problems with bounded controls (Ref. 8).

In the area of ordinary gradient methods, Miele et al (Ref. 9) developed a sequential gradient-restoration algorithm for optimal control problems where the state $x(t)$, the control $u(t)$, and the parameter π must satisfy not only differential constraints and terminal constraints, but also nondifferential constraints everywhere along the interval of integration. The importance of Ref. 9 lies in that (i) many optimization problems arise directly in the form considered there, (ii) problems involving equality constraints can be reduced to that form through suitable transformations, and (iii) problems involving inequality constraints can be reduced to that form through suitable transformations. Thus, an extremely large class of problems can be handled. This includes problems with bounded control, bounded state, bounded time rate of change of the state, as well as problems where a bound is imposed on some function of the parameter, the control, the state, and the time rate of change of the state.

This report combines the ideas of Ref. 6 and those of Ref. 9. The result is a sequential conjugate gradient-restoration algorithm which can handle constrained minimization problems, characterized by the presence of nondifferential constraints, without resorting to penalty functions. The algorithm is composed of a

sequence of cycles, each cycle consisting of two phases, a conjugate gradient phase and a restoration phase.

The conjugate gradient phase involves a single iteration and is designed to decrease the value of the functional while satisfying the constraints to first order. During this iteration, the first variation of the functional is minimized subject to the linearized constraints. The minimization is performed over the class of variations of the control and the parameter which are equidistant from some constant multiple of the corresponding variations of the previous conjugate gradient phase.

The restoration phase involves one or more iterations and is designed to restore the constraints to a predetermined accuracy while the norm of the variations of the control and the parameter is minimized, subject to the linearized constraints.

The sequential conjugate gradient-restoration algorithm is characterized by two main properties. First, at the end of each conjugate gradient-restoration cycle, the trajectory satisfies the constraints to a given accuracy; thus, a sequence of feasible suboptimal solutions is produced. Second, the conjugate gradient stepsize and the restoration stepsize can be chosen so that the restoration phase preserves the descent property of the conjugate gradient phase; thus, the value of the functional at the end of any cycle is smaller than the value of the functional at the beginning of that cycle. Of course,

restarting the algorithm might be occasionally necessary. For a discussion of the basic properties of the sequential gradient-restoration algorithm, see Ref. 10.

2. Formulation of the Problem

Consider the problem of minimizing the functional

$$I = \int_0^\tau f_*(x,u,\pi_*,\theta)\,d\theta + [g_*(x,\pi_*,\theta)]_\tau \qquad (1)$$

with respect to the state $x(\theta)$, the control $u(\theta)$, and the parameters π_* and τ which satisfy the differential constraint

$$dx/d\theta = \phi_*(x,u,\pi_*,\theta), \qquad 0 \leq \theta \leq \tau, \qquad (2)$$

the nondifferential constraint

$$S_*(x,u,\pi_*,\theta) = 0, \qquad 0 \leq \theta \leq \tau, \qquad (3)$$

and the boundary conditions

$$(x)_0 = \text{given}, \qquad [\psi_*(x,\pi_*,\theta)]_\tau = 0. \qquad (4)$$

In the above equations, the functions f_* and g_* are scalar, the function ϕ_* is an n-vector, the function S_* is a k-vector, and the function ψ_* is a q-vector. The independent variable is the actual time θ, while the dependent variables are the state x (an n-vector), the control u (an m-vector), the parameter π_* (a p_*-vector), and the parameter τ(a scalar).

At the initial time $\theta = 0$, the n components of the vector x are specified. At the final time $\theta = \tau$, q scalar relations are specified, where $q \leq n + p_*$ if τ is fixed and $q \leq n + p_* + 1$ if τ is free.

To facilitate the implementation of the algorithm on a digital computer, we replace the actual time θ with the normalized time t. The latter is defined in such a way that the interval of integration has unit length. Thus, in normalized form, $t = 0$ denotes the time at which the initial boundary (4-1) is left and $t = 1$ denotes the time at which the terminal boundary (4-2) is reached. The following linear relation allows the passage from the normalized time t to the actual time θ:

$$\theta = \tau t, \qquad 0 \leq t \leq 1. \tag{5}$$

The fact that the normalized final time is fixed (t=1) does not cause any loss of generality in the problem. If the actual final time is free, τ simply becomes a parameter to be optimized in the transformed problem. In view of this, we define the augmented parameter π (a p-vector),

$$\pi = \pi_* \text{ or } \pi = \begin{bmatrix} \pi_* \\ \tau \end{bmatrix}, \tag{6}$$

where (6-1) holds if τ is fixed and (6-2) holds if τ is free.

In addition to the normalized time t and the parameter π,

we define the following functions:

$$f(x,u,\pi,t) = \tau f_*(x,u,\pi_*,\tau t), \qquad 0 \le t \le 1, \tag{7}$$

$$\phi(x,u,\pi,t) = \tau\phi_*(x,u,\pi_*,\tau t), \qquad 0 \le t \le 1, \tag{8}$$

$$S(x,u,\pi,t) = S_*(x,u,\pi_*,\tau t), \qquad 0 \le t \le 1, \tag{9}$$

$$g(x,\pi,t) = g_*(x,\pi_*,\tau t), \tag{10}$$

$$\psi(x,\pi,t) = \psi_*(x,\pi_*,\tau t). \tag{11}$$

Under the above transformations and definitions, problem (1)-(4) can be reformulated as follows. Minimize the functional

$$I = \int_0^1 f(x,u,\pi,t)dt + [g(x,\pi,t)]_1 \tag{12}$$

with respect to the state $x(t)$, the control $u(t)$, and the parameter π which satisfy the differential constraint

$$\dot{x} = \phi(x,u,\pi,t), \qquad 0 \le t \le 1, \tag{13}$$

the nondifferential constraint

$$S(x,u,\pi,t) = 0, \qquad 0 \le t \le 1, \tag{14}$$

and the boundary conditions

$$(x)_0 = \text{given}, \qquad [\psi(x,\pi,t)]_1 = 0. \tag{15}$$

From calculus of variations, we know that the problem (12)-(15) is one of the Bolza type; it can be recast as that of minimizing the augmented functional[5]

$$J = \int_0^1 (\lambda^T \dot{x} + H)\, dt + (G)_1$$

$$= (\lambda^T x)_0^1 + \int_0^1 (H - \dot{\lambda}^T x)\, dt + (G)_1 \qquad (16)$$

with respect to the state x(t), the control u(t), and the parameter π which satisfy (13)-(15), where the functions H and G are given by

$$H = f - \lambda^T \phi + \rho^T S, \qquad G = g + \mu^T \psi, \qquad (17)$$

and where $\lambda(t)$ is a variable Lagrange multiplier (an n-vector), $\rho(t)$ is a variable Lagrange multiplier (a k-vector), and μ is a constant Lagrange multiplier (a q-vector).

The optimal solution must satisfy (13)-(15) and the first-order optimality conditions, namely, the Euler equations

$$\dot{\lambda} = H_x, \qquad H_u = 0, \qquad \int_0^1 H_\pi dt + (G_\pi)_1 = 0 \qquad (18)$$

---

[5] In Eq. (16), it is tacitly assumed that the initial condition (15-1) is satisfied. The second form of Eq. (16) arises after the customary integration by parts is performed.

and the following natural condition arising from the transversality condition:

$$(\lambda + G_x)_1 = 0. \tag{19}$$

Summarizing, we seek functions $x(t)$, $u(t)$, π and multipliers $\lambda(t)$, $\rho(t)$, μ which satisfy the constraints (13)-(15) and the optimality conditions (20)-(23):

$$\dot{\lambda} - f_x + \phi_x\lambda - S_x\rho = 0, \qquad 0 \le t \le 1, \tag{20}$$

$$f_u - \phi_u\lambda + S_u\rho = 0, \qquad 0 \le t \le 1, \tag{21}$$

$$\int_0^1 (f_\pi - \phi_\pi\lambda + S_\pi\rho)dt + (g_\pi + \psi_\pi\mu)_1 = 0 , \tag{22}$$

$$(\lambda + g_x + \psi_x\mu)_1 = 0. \tag{23}$$

2.1. Approximate Methods. Since the differential system (13)-(15) and (20)-(23) is generally nonlinear, some iterative technique must be employed in its solution. For this purpose, let us define the scalar functionals P and Q, which denote the constraint error and the error in the optimality conditions, respectively. We have

$$P = \int_0^1 N(\dot{x} - \phi)dt + \int_0^1 N(S)dt + N(\psi)_1 , \tag{24}$$

$$Q = \int_0^1 N(\dot{\lambda} - f_x + \phi_x \lambda - S_x \rho)\, dt + \int_0^1 N(f_u - \phi_u \lambda + S_u \rho)\, dt$$

$$+ N\left[\int_0^1 (f_\pi - \phi_\pi \lambda + S_\pi \rho)\, dt + (g_\pi + \psi_\pi \mu)_1\right] + N(\lambda + g_x + \psi_x \mu)_1, \quad (25)$$

where $N(b)$ denotes the norm squared of a vector, i.e.,

$$N(b) = b^T b \quad (26)$$

for a given vector b.

Note that, for the optimal solution, $P = 0$ and $Q = 0$. For an approximation to the optimal solution,

$$P \leq \varepsilon_1, \qquad Q \leq \varepsilon_2, \quad (27)$$

where ε_1 and ε_2 are small, preselected numbers.

3. Construction of the Sequential Conjugate Gradient-Restoration Algorithm.

The sequential conjugate gradient-restoration algorithm is an iterative technique which includes a sequence of cycles having the following properties.[6]

Property 3.1. The functions $x(t)$, $u(t)$, π available both at the beginning and at the end of each cycle must be feasible; that is, they must be consistent with the constraints (13)-(15) within the preselected accuracy (27-1).

Property 3.2. The functions $x(t)$, $u(t)$, π produced at the end of each cycle must be characterized by a value of the functional I [see Eq. (12)] which is smaller than that associated with the functions available at the beginning of the cycle.

Property 3.3. The functions $x(t)$, $u(t)$, π produced at the end of each cycle must be characterized by a value of the augmented functional J [see Eq. (16)] which is smaller than that associated with the functions available at the beginning of the cycle.

---

[6] Note that Property 3.3 is a consequence of Properties 3.1 and 3.2. Conversely, Property 3.2 is a consequence of Properties 3.1 and 3.3.

To achieve the above properties, each cycle is made of two phases, a conjugate gradient phase and a restoration phase.

Conjugate Gradient Phase. This phase is started only when the constraint error P satisfies Ineq. (27-1). It involves a single iteration, which is designed to decrease the value of the functional I or the augmented functional J, while satisfying the constraints to first order. During this iteration, the first variation of the functional I is minimized, subject to the linearized constraints. The minimization is performed over the class of variations of the control and the parameter which are equidistant from some constant multiple of the corresponding variations of the previous conjugate gradient phase.

Restoration Phase. This phase is started only when the constraint error P violates Ineq. (27-1). The restoration phase involves one or more iterations. In each restorative iteration, the objective is to reduce the constraint error P, while the constraints are satisfied to first order and the norm of the variations of the control and the parameter is minimized. The restoration phase is terminated whenever Ineq. (27-1) is satisfied.

Remark. During each conjugate gradient iteration or restorative iteration, use of nonlinear equations must be avoided. Therefore, the exact feasibility equations (13)-(15) are replaced by their corresponding linearized approximations. These

linearized approximations do not include forcing terms in the conjugate gradient phase, while they do include forcing terms in the restoration phase.

Notation. For any iteration of the conjugate gradient phase or the restoration phase, the following terminology is adopted: $x(t)$, $u(t)$, π denote the nominal functions; $\tilde{x}(t)$, $\tilde{u}(t)$, $\tilde{\pi}$ denote the varied functions; and $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$ denote the displacements leading from the nominal functions to the varied functions. These quantities satisfy the definitions

$$\tilde{x}(t) = x(t) + \Delta x(t), \qquad \tilde{u}(t) = u(t) + \Delta u(t), \qquad \tilde{\pi} = \pi + \Delta\pi. \tag{28}$$

If the variations appearing in (28) are linear in the stepsize α, where $\alpha > 0$, they take the form

$$\Delta x(t) = \alpha A(t), \qquad \Delta u(t) = \alpha B(t), \qquad \Delta\pi = \alpha C, \tag{29}$$

with the implication that

$$\tilde{x}(t) = x(t) + \alpha A(t), \qquad \tilde{u}(t) = u(t) + \alpha B(t), \qquad \tilde{\pi} = \pi + \alpha C. \tag{30}$$

The functions $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$ must be determined so as to produce some desirable effect at every iteration, namely, the decrease of the functionals I and/or J and/or P. Thus, the

following descent properties are required:

$$\tilde{I} < I , \quad \text{and/or} \quad \tilde{J} < J , \quad \text{and/or} \quad \tilde{P} < P, \tag{31}$$

where I,J,P are associated with the nominal functions and $\tilde{I}, \tilde{J}$, $\tilde{P}$ are associated with the varied functions. Inequalities (31-1) and (31-2) characterize the conjugate gradient phase, and Ineq. (31-3) characterizes the restoration phase.

In turn, relations (31) can be enforced at every iteration providing the stepsize α is sufficiently small and the functions A(t), B(t), C are chosen so that

$$\delta I < 0, \quad \text{and/or} \quad \delta J < 0, \quad \text{and/or} \quad \delta P < 0, \tag{32}$$

where the symbol $\delta(\ldots)$ denotes the first variation. Inequalities (32-1) and (32-2) characterize the conjugate gradient phase, and Ineq. (32-3) characterizes the restoration phase.

Clearly, every iteration includes two basic operations: (a) the determination of functions A(t), B(t), C consistent with the first variation requirements (32); and (b) the determination of the stepsize α consistent with the total variation requirements (31).

Outline. In Section 4, we describe the equations of the restoration phase; we show how nominal functions consistent with the feasibility equations (13)-(15) can be obtained. In

Section 5, we describe the general equations of the conjugate gradient phase; the linear case (case where the constraints are linear) is treated in Section 6; the linear-quadratic case (case where the functional is quadratic and the constraints are linear) is treated in Section 7; here, we show that certain general conjugacy and orthogonality conditions hold. Always with reference to the conjugate gradient phase, the nonlinear-nonquadratic case (case where the functional is nonquadratic and/or the constraints are nonlinear) is treated in Section 8; here, we discuss the implementation of a first-order algorithm (this is an algorithm which uses first derivatives at most). In Section 9, we discuss the descent property of a complete conjugate gradient-restoration cycle. In Section 10, we present a summary of the sequential conjugate gradient-restoration algorithm. Finally, in Section 11, we list the safeguards necessary to its implementation on a digital computer.

4. Restoration Phase

As stated before, the restoration phase is started only when the constraint error P violates Ineq. (27-1). The restoration phase involves one or more iterations. In each restorative iteration, the objective is to reduce the constraint error P, while the constraints are satisfied to first order and the norm of the variations of the control and the parameter is minimized. The restoration phase is terminated whenever Ineq. (27-1) is satisfied.

There are two situations where the restoration phase is employed: (a) at the very beginning of the algorithm, one needs to generate nominal functions consistent with the feasibility equations (13)-(15); and (b) subsequently, one needs to correct possible constraint violations occurring during a conjugate gradient phase: these constraint violations are due to the fact that Eqs. (13)-(15) are considered only in linearized form during a conjugate gradient phase.

Linearized Equations. Let $x(t)$, $u(t)$, π denote nominal functions not satisfying (13)-(15), and let $\tilde{x}(t)$, $\tilde{u}(t)$, $\tilde{\pi}$ denote varied functions satisfying (13)-(15). To first order, the perturbations $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$ must satisfy the linearized constraint equations

$$\Delta\dot{x} - \phi_x^T \Delta x - \phi_u^T \Delta u - \phi_\pi^T \Delta\pi + \alpha(\dot{x} - \phi) = 0, \qquad 0 \le t \le 1, \tag{33}$$

$$S_x^T \Delta x + S_u^T \Delta u + S_\pi^T \Delta \pi + \alpha S = 0, \qquad 0 \le t \le 1, \tag{34}$$

$$(\Delta x)_0 = 0, \qquad (\psi_x^T \Delta x + \psi_\pi^T \Delta \pi + \alpha \psi)_1 = 0, \tag{35}$$

where α denotes a scaling factor (restoration stepsize) in the range $0 \le \alpha \le 1$.

Descent Property. The linearized equations (33)-(35) admit an infinite number of solutions. Each of these solutions is characterized by a descent property in the constraint error P. This can be seen by computing the first variation of the functional (24):

$$\delta P = 2\int_0^1 (\dot{x} - \phi)^T (\Delta \dot{x} - \phi_x^T \Delta x - \phi_u^T \Delta u - \phi_\pi^T \Delta \pi)\, dt$$
$$+ 2\int_0^1 S^T (S_x^T \Delta x + S_u^T \Delta u + S_\pi^T \Delta \pi)\, dt + 2[\psi^T (\psi_x^T \Delta x + \psi_\pi^T \Delta \pi)]_1 \tag{36}$$

and by observing that, when the perturbations defined by (33)-(35) are employed, the first variation of the constraint error (36) becomes

$$\delta P = -2\alpha P. \tag{37}$$

Since $P > 0$, Eq. (37) shows that $\delta P < 0$. Hence, for α sufficiently small, a decrease in the constraint error P is guaranteed.

Auxiliary Minimization Problem. Since Eqs. (33)-(35) admit an infinite number of solutions, an additional requirement must be introduced in order to uniquely define the restorative iteration. More specifically, we require that restoration be accomplished with the least-square change of the control and the parameter (Ref. 10). Thus, we seek the minimum of the quadratic functional

$$K = (1/2\alpha)\left[\int_0^1 \Delta u^T \Delta u dt + \Delta\pi^T \Delta\pi\right] \tag{38}$$

with respect to the perturbations $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$ which satisfy the linearized constraints (33)-(35).

Special Variations. From calculus of variations, we know that the problem represented by (33),(34),(35),(38) is one of the Bolza type. In this connection, let $\lambda(t)$ denote a variable Lagrange multiplier associated with the differential constraint (33); let $\rho(t)$ denote a variable Lagrange multiplier associated with the nondifferential constraint (34); and let μ denote a constant Lagrange multiplier associated with the final condition (35-2). With this understanding, the Euler equations optimizing $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$ and the natural condition arising from the transversality condition are written as

$$\dot{\lambda} + \phi_x \lambda - S_x \rho = 0, \qquad 0 \le t \le 1, \tag{39}$$

$$\Delta u = \alpha(\phi_u \lambda - S_u \rho), \qquad 0 \le t \le 1, \tag{40}$$

$$\Delta\pi = \alpha\left[\int_0^1 (\phi_\pi \lambda - S_\pi \rho)\,dt - (\psi_\pi \mu)_1\right], \tag{41}$$

$$(\lambda + \psi_x \mu)_1 = 0. \tag{42}$$

Summarizing, we seek variations $\Delta x(t), \Delta u(t), \Delta\pi$ and multipliers $\lambda(t), \rho(t), \mu$ which satisfy the constraints (33)-(35) and the optimality conditions (39)-(42).

Basic Functions. If the definitions (29) are invoked, the stepsize α can be eliminated, and Eqs. (33)-(35) and (39)-(42) can be rewritten as

$$\dot{A} - \phi_x^T A - \phi_u^T B - \phi_\pi^T C + (\dot{x} - \phi) = 0, \qquad 0 \le t \le 1, \tag{43}$$

$$S_x^T A + S_u^T B + S_\pi^T C + S = 0, \qquad 0 \le t \le 1, \tag{44}$$

$$(A)_0 = 0, \qquad (\psi_x^T A + \psi_\pi^T C + \psi)_1 = 0, \tag{45}$$

$$\dot{\lambda} + \phi_x \lambda - S_x \rho = 0, \qquad 0 \le t \le 1, \tag{46}$$

$$B - \phi_u \lambda + S_u \rho = 0, \qquad 0 \le t \le 1, \tag{47}$$

$$C + \int_0^1 (-\phi_\pi \lambda + S_\pi \rho)\,dt + (\psi_\pi \mu)_1 = 0, \tag{48}$$

$$(\lambda + \psi_x \mu)_1 = 0. \tag{49}$$

Equations (43)-(49) uniquely define the basic functions A(t),

B(t), C as well as the multipliers $\lambda(t)$, $\rho(t)$, μ of the restoration phase.

Solution Technique. Let y denote the $(n+p)$-vector

$$y = \begin{bmatrix} \lambda(0) \\ C \end{bmatrix}. \tag{50}$$

Let a sweep be defined as a forward integration of the system (43)-(49) obtained by (a) assigning a particular value to the vector y, (b) employing Eqs. (43), (44), (45-1), (46), (47), and (c) bypassing Eqs. (45-2), (48), (49).

Let $n+p+1$ independent sweeps be executed. More specifically, the first $n+p$ sweeps are executed by choosing the $n+p$ linearly independent vectors y to be the columns of the identity matrix of order $n+p$. The last sweep is executed by choosing y as the null vector. In this way, we obtain the particular solutions (Refs. 11-14)

$$A_i(t),\ B_i(t),\ C_i,\ \lambda_i(t),\ \rho_i(t), \qquad i = 1,2,\ldots,n+p+1. \tag{51}$$

Then, we introduce the $n+p+1$ undetermined, scalar constant k_i and form the linear combinations

$$A(t) = \Sigma k_i A_i(t), \qquad B(t) = \Sigma k_i B_i(t), \qquad C = \Sigma k_i C_i, \tag{52}$$

$$\lambda(t) = \Sigma k_i \lambda_i(t), \qquad \rho(t) = \Sigma k_i \rho_i(t), \tag{53}$$

where the summation is taken over the index i. The $n+p+1$ coefficients k_i and the q components of the vector μ are obtained by forcing the linear combinations (52)-(53) to satisfy Eqs. (45-2), (48), (49), together with the normalization condition (Refs. 11-14)

$$\Sigma k_i = 1. \tag{54}$$

Stepsize. With the basic functions A(t), B(t), C known, we consider the one-parameter family of solutions (30). For this one-parameter family, the constraint error (24) becomes a function of the form

$$\tilde{P} = \tilde{P}(\alpha) \quad . \tag{55}$$

Then, the stepsize α must be selected so that the inequality

$$\tilde{P}(\alpha) < \tilde{P}(0) \tag{56}$$

is satisfied while keeping

$$\tilde{\tau}(\alpha) \geq 0 \quad . \tag{57}$$

Satisfaction of Ineqs. (56) and (57) is guaranteed for α sufficiently small. Any violation of the above inequalities

necessitates a reduction in the stepsize. Such a reduction can be obtained by employing a bisection process, starting from the reference stepsize

$$\alpha_0 = 1 \quad . \tag{58}$$

This reference stepsize has the following property: it yields one-step restoration for the limiting case where the constraints (13)-(15) are linear.

5. Conjugate Gradient Phase: General Case

As stated before, the conjugate gradient phase is started only when the constraint error P satisfies Ineq. (27-1). It involves a single iteration, which is designed to decrease the value of the functional I or the augmented functional J, while satisfying the constraints to first order. During this iteration, the first variation of the functional I is minimized, subject to the linearized constraints. The minimization is performed over the class of variations of the control and the parameter which are equidistant from some constant multiple of the corresponding variations of the previous conjugate gradient phase.

For the sake of clarity, the general structure of the conjugate gradient phase is given first in this section. The linear case is treated in Section 6, and the linear-quadratic case is treated in Section 7. Then, the extension to the nonlinear-nonquadratic case is given in Section 8.

Linearized Equations. Let $x(t)$, $u(t)$, π denote nominal functions satisfying (13)-(15),[7] and let $\tilde{x}(t)$, $\tilde{u}(t)$, $\tilde{\pi}$ denote

[7] These nominal functions can be obtained by employing the restoration algorithm of Section 4.

varied functions also satisfying (13)-(15). To first order, the perturbations $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$ must satisfy the linearized constraint equations

$$\Delta\dot{x} - \phi_x^T \Delta x - \phi_u^T \Delta u - \phi_\pi^T \Delta\pi = 0, \qquad 0 \le t \le 1, \tag{59}$$

$$S_x^T \Delta x + S_u^T \Delta u + S_\pi^T \Delta\pi = 0, \qquad 0 \le t \le 1, \tag{60}$$

$$(\Delta x)_0 = 0, \qquad (\psi_x^T \Delta x + \psi_\pi^T \Delta\pi)_1 = 0 . \tag{61}$$

Note the difference between Eqs. (33)-(35) and Eqs. (59)-(61). While the former are nonhomogeneous, the latter are homogeneous in the perturbations $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$.

Auxiliary Minimization Problem. Since the linearized equations (59)-(61) admit an infinite number of solutions, some additional requirement must be introduced in order to uniquely define the conjugate gradient iteration. More specifically, we consider the first variation of the functional (12):

$$\delta I = \int_0^1 (f_x^T \Delta x + f_u^T \Delta u + f_\pi^T \Delta\pi)dt + (g_x^T \Delta x + g_\pi^T \Delta\pi)_1 \tag{62}$$

and the isoperimetric constraint

$$K = \int_0^1 (\Delta u - \beta\Delta\hat{u})^T (\Delta u - \beta\Delta\hat{u})dt + (\Delta\pi - \beta\Delta\hat{\pi})^T (\Delta\pi - \beta\Delta\hat{\pi}), \tag{63}$$

where K and β are undetermined constants. The symbols $\Delta\hat{x}(t)$, $\Delta\hat{u}(t)$, $\Delta\hat{\pi}$ denote the variations associated with the previous conjugate gradient iteration. Therefore, because of (29), we have

$$\Delta\hat{x}(t) = \hat{\alpha}\,\hat{A}(t), \qquad \Delta\hat{u}(t) = \hat{\alpha}\,\hat{B}(t), \qquad \Delta\hat{\pi} = \hat{\alpha}\,\hat{C}\,. \tag{64}$$

Then, we seek the minimum of the linear functional (62) with respect to the perturbations $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$ which satisfy the linearized constraints (59)-(61) and the quadratic isoperimetric constraint (63).

Special Variations. From calculus of variations, we know that the problem represented by (59)-(63) is one of the Bolza type with an added isoperimetric constraint on the variations of the control and the parameter. In this connection, let $\lambda(t)$ denote a variable Lagrange multiplier associated with the differential constraint (59); let $\rho(t)$ denote a variable Lagrange multiplier associated with the nondifferential constraint (60); let μ denote a constant Lagrange multiplier associated with the final condition (61-2); and let $1/2\alpha$ denote a constant Lagrange multiplier associated with the isoperimetric constraint (63). With this understanding, the Euler equations optimizing $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$ and the natural condition arising from the transversality condition are written as

$$\dot{\lambda} - f_x + \phi_x\lambda - S_x\rho = 0, \qquad 0 \le t \le 1, \tag{65}$$

$$(\Delta u - \beta\Delta\hat{u})/\alpha + f_u - \phi_u\lambda + S_u\rho = 0, \qquad 0 \le t \le 1, \tag{66}$$

$$(\Delta\pi - \beta\Delta\hat{\pi})/\alpha + \int_0^1 (f_\pi - \phi_\pi\lambda + S_\pi\rho)dt + (g_\pi + \psi_\pi\mu)_1 = 0, \tag{67}$$

$$(\lambda + g_x + \psi_x\mu)_1 = 0 \quad . \tag{68}$$

Summarizing, we seek variations $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$ and multipliers $\lambda(t)$, $\rho(t)$, μ, $1/2\alpha$ which satisfy the constraints (59), (60), (61), (63) and the optimality conditions (65)-(68).

Basic Functions. Let the definitions (29) be invoked for both the present conjugate gradient phase and the previous conjugate gradient phase. Let the directional coefficient γ be defined as

$$\gamma = \beta(\hat{\alpha}/\alpha) \quad . \tag{69}$$

With this understanding, the stepsize α can be eliminated, and Eqs. (59)-(61) and (65)-(68) can be rewritten as

$$\dot{A} - \phi_x^T A - \phi_u^T B - \phi_\pi^T C = 0, \qquad 0 \le t \le 1, \tag{70}$$

$$S_x^T A + S_u^T B + S_\pi^T C = 0, \qquad 0 \le t \le 1, \tag{71}$$

$$(A)_0 = 0, \qquad (\psi_x^T A + \psi_\pi^T C)_1 = 0, \tag{72}$$

$$\dot{\lambda} - f_x + \phi_x \lambda - S_x \rho = 0, \qquad 0 \leq t \leq 1, \tag{73}$$

$$B - \gamma \hat{B} + f_u - \phi_u \lambda + S_u \rho = 0, \qquad 0 \leq t \leq 1, \tag{74}$$

$$C - \gamma \hat{C} + \int_0^1 (f_\pi - \phi_\pi \lambda + S_\pi \rho) dt + (g_\pi + \psi_\pi \mu)_1 = 0, \tag{75}$$

$$(\lambda + g_x + \psi_x \mu)_1 = 0 . \tag{76}$$

For a given value of the directional coefficient γ, Eqs. (70)-(76) uniquely define the basic functions $A(t)$, $B(t)$, C as well as the multipliers $\lambda(t)$, $\rho(t)$, μ of the conjugate gradient phase.

Isoperimetric Constant. In the light of (29) and (69), the isoperimetric functional (63) takes the form

$$K = \alpha^2 \left[\int_0^1 (B - \gamma \hat{B})^T (B - \gamma \hat{B}) dt + (C - \gamma \hat{C})^T (C - \gamma \hat{C}) \right] . \tag{77}$$

If the basic functions $A(t)$, $B(t)$, C are consistent with (70)-(76), the error in the optimality conditions (25) reduces to

$$Q = \int_0^1 (B - \gamma \hat{B})^T (B - \gamma \hat{B}) dt + (C - \gamma \hat{C})^T (C - \gamma \hat{C}) . \tag{78}$$

Consequently, the following relation ties the isoperimetric constant, the stepsize, and the error in the optimality conditions:

$$K = \alpha^2 Q . \tag{79}$$

Clearly, to assign values to the isoperimetric constant is the same as assigning values to the stepsize. However, there is no clear-cut way of determining a priori convenient values for the isoperimetric constant. Therefore, the implementation of the conjugate gradient algorithm becomes simpler if one avoids evaluating α in terms of K through (79) and assigns values to α directly.

Descent Property. Next, consider the augmented functional (16) and its first variation

$$\delta J = \int_0^1 (f_x - \phi_x\lambda + S_x\rho - \dot{\lambda})^T \Delta x dt + \int_0^1 (f_u - \phi_u\lambda + S_u\rho)^T \Delta u dt$$
$$+ \left[\int_0^1 (f_\pi - \phi_\pi\lambda + S_\pi\rho) dt + (g_\pi + \psi_\pi\mu)_1\right]^T \Delta\pi + \left[(\lambda + g_x + \psi_x\mu)^T \Delta x\right]_1 . \quad (80)$$

When the perturbations defined by (29) and (70)-(76) are employed, Eq. (80) becomes

$$\delta J = -\alpha\left[\int_0^1 (B - \gamma\hat{B})^T B dt + (C - \gamma\hat{C})^T C\right] . \quad (81)$$

Upon invoking Eq. (78) and defining the quantity

$$Z = \int_0^1 (B - \gamma B)^T \hat{B} dt + (C - \gamma\hat{C})^T \hat{C} , \quad (82)$$

we see that Eq. (81) can be rewritten as

$$\delta J = -\alpha(Q + \gamma Z) \quad . \tag{83}$$

For the first iteration of the conjugate gradient phase, one sets

$$\gamma = 0, \tag{84}$$

with the implication that

$$\delta J = -\alpha Q \quad . \tag{85}$$

Since $Q > 0$, Eq. (85) shows that $\delta J<0$. Hence, for α sufficiently small, a decrease in the augmented functional J is guaranteed.

For subsequent iterations, one sets $\gamma \neq 0$. More specifically, the directional coefficient must be such that

$$\gamma > 0, \tag{86}$$

and its proper value is discussed in Section 7. At any rate, Eq. (83) shows that $\delta J<0$ providing

$$Q + \gamma Z > 0 \ , \tag{87}$$

where Q is given by (78) and Z is given by (82). Hence, for

α sufficiently small, the decrease in the augmented functional J is guaranteed as long as Ineq. (87) is satisfied. If Ineq. (87) is violated, the descent property on J no longer holds, and the conjugate gradient phase must be restarted by resetting the directional coefficient γ at the level (84).

Solution Technique. Now, assume that a particular value is given to the directional coefficient γ. Let y denote the $(n + p)$-vector

$$y = \begin{bmatrix} \lambda(0) \\ C \end{bmatrix}. \tag{88}$$

Let a sweep be defined as a forward integration of the system (70)-(76) obtained by (a) assigning a particular value to the vector y, (b) employing Eqs. (70), (71), (72-1), (73), (74), and (c) bypassing Eqs. (72-2), (75), (76).

Let $n + p + 1$ independent sweeps be executed. More specifically, the first $n + p$ sweeps are executed by choosing the $n + p$ linearly independent vectors y to be the columns of the identity matrix of order $n + p$. The last sweep is executed by choosing y as the null vector. In this way, we obtain the particular solutions (Refs. 11-14)

$$A_i(t),\ B_i(t),\ C_i,\ \lambda_i(t),\ \rho_i(t), \qquad i = 1, 2, \ldots, n + p + 1. \tag{89}$$

Then, we introduce the $n+p+1$ undetermined, scalar constants k_i and form the linear combinations

$$A(t) = \Sigma k_i A_i(t), \qquad B(t) = \Sigma k_i B_i(t), \qquad C = \Sigma k_i C_i, \tag{90}$$

$$\lambda(t) = \Sigma k_i \lambda_i(t), \qquad \rho(t) = \Sigma k_i \rho_i(t), \tag{91}$$

where the summation is taken over the index i. The $n+p+1$ coefficients k_i and the q components of the vector μ are obtained by forcing the linear combinations (90)-(91) to satisfy Eqs. (72-2), (75), (76), together with the normalization condition (Refs. 11-14)

$$\Sigma k_i = 1 \ . \tag{92}$$

General Solution. Next, assume that two particular values are given to the directional coefficient γ, for instance,

$$\gamma_* = 0 \qquad \text{and} \qquad \gamma_{**} = 1 \ . \tag{93}$$

Assume that the previous solution technique is employed twice, and denote by

$$A_*(t),\ B_*(t),\ C_*,\ \lambda_*(t),\ \rho_*(t),\ \mu_* \tag{94}$$

and

$$A_{**}(t),\ B_{**}(t),\ C_{**},\ \lambda_{**}(t),\ \rho_{**}(t),\ \mu_{**} \tag{95}$$

the particular solutions of (70)-(76) corresponding to (93-1) and (93-2), respectively. Simple manipulations, omitted for the sake of brevity, show that the general solution of (70)-(76), valid for any value of the directional coefficient γ, can be written as

$$A(t) = A_*(t) + \gamma[A_{**}(t) - A_*(t)], \qquad B(t) = B_*(t) + \gamma[B_{**}(t) - B_*(t)],$$

$$C = C_* + \gamma(C_{**} - C_*), \tag{96}$$

$$\lambda(t) = \lambda_*(t) + \gamma[\lambda_{**}(t) - \lambda_*(t)], \qquad \rho(t) = \rho_*(t) + \gamma[\rho_{**}(t) - \rho_*(t)],$$

$$\mu = \mu_* + \gamma(\mu_{**} - \mu_*)\ . \tag{97}$$

As a conclusion, the general solution of (70)-(76) requires that two sweeps of $n+p+1$ integrations be executed, one leading to the particular solution (94) and one leading to the particular solution (95). However, if the constraints are linear, the general solution of (70)-(76) requires only one sweep of $n+p+1$ integrations, that leading to the particular solution (94), as is shown in Section 6.

Stepsize and Directional Coefficient. After the general solution (96)-(97) is known, the next step is to determine the

proper values of the stepsize α and the directional coefficient γ. A logical scheme is that of determining these quantities so that the augmented functional (16) is minimized.

For the varied functions $\tilde{x}(t)$, $\tilde{u}(t)$, $\tilde{\pi}$, let the augmented functional (16) be written in the form

$$\tilde{J} = (\lambda^T\tilde{x})_0^1 + \int_0^1 (\tilde{f} - \lambda^T\tilde{\phi} + \rho^T\tilde{S} - \dot{\lambda}^T\tilde{x})dt + (\tilde{g} + \mu^T\tilde{\psi})_1 \ . \tag{98}$$

In view of (30) and (96), the varied functions $\tilde{x}(t)$, $\tilde{u}(t)$, $\tilde{\pi}$ constitute the two-parameter family

$$\tilde{x}(t) = x(t) + \alpha \left\{A_*(t) + \gamma[A_{**}(t) - A_*(t)]\right\} \ , \tag{99-1}$$

$$\tilde{u}(t) = u(t) + \alpha \left\{B_*(t) + \gamma[B_{**}(t) - B_*(t)]\right\} \ , \tag{99-2}$$

$$\tilde{\pi} = \pi + \alpha[C_* + \gamma(C_{**} - C_*)] \ . \tag{99-3}$$

On the other hand, the multipliers $\lambda(t)$, $\rho(t)$, μ constitute the one-parameter family (97). Upon using (97) and (99), we see that the augmented functional (98) takes the form

$$\tilde{J} = \tilde{J}(\alpha, \gamma) \ . \tag{100}$$

Therefore, the optimum values of α and γ satisfy the relations

$$\tilde{J}_\alpha(\alpha,\gamma) = 0, \qquad \tilde{J}_\gamma(\alpha,\gamma) = 0 \ . \tag{101}$$

In principle, one can solve Eqs. (101) using an exact search, as done in Ref. 15 for mathematical programming problems. The resulting algorithm constitutes the extension to optimal control problems of the memory-gradient method of Ref. 15. However, the simultaneous determination of α and γ might be expensive computationally; this being the case, we follow a different road. First, we determine an approximate value of the directional coefficient γ, based on the consideration of the linear-quadratic case (Section 7). Once γ is known, the two-parameter family (100) reduces to the one-parameter family

$$\tilde{J} = \tilde{J}(\alpha) \ . \tag{102}$$

Then, the optimum stepsize α satisfies the relation

$$\tilde{J}_\alpha(\alpha) = 0, \tag{103}$$

whose numerical solution can be arrived at in a variety of ways. For example, within the frame of the linear-quadratic case, the numerical solution of (103) can be obtained with quadratic interpolation (Section 8). On the other hand, within the frame of the nonlinear-nonquadratic case, the numerical

solution of (103) can be obtained with cubic interpolation (Section 8).

6. Conjugate Gradient Phase: Linear Constraints

In the previous section, we derived some general relations which are valid for the conjugate gradient phase, regardless of the analytical form of the functional (12) and the constraints (13)-(15). In this section, we give the particular relations which are valid if the constraints are linear.

General Solution. Under the linearity assumption for the constraints, consider the system (70)-(76) which defines the basic functions $A(t)$, $B(t)$, C as well as the multipliers $\lambda(t)$, $\rho(t)$, μ. By substitution, it can be verified that the particular solutions (94) and (95) satisfy the relations

$$A_{**}(t) - A_*(t) = \hat{A}(t), \qquad B_{**}(t) - B_*(t) = \hat{B}(t), \qquad C_{**} - C_* = \hat{C}, \tag{104}$$

$$\lambda_{**}(t) - \lambda_*(t) = 0, \qquad \rho_{**}(t) - \rho_*(t) = 0, \qquad \mu_{**} - \mu_* = 0 . \tag{105}$$

As a consequence, Eqs. (96)-(97) take the simpler form

$$A(t) = A_*(t) + \gamma\hat{A}(t), \qquad B(t) = B_*(t) + \gamma\hat{B}(t), \qquad C = C_* + \gamma\hat{C}, \tag{106}$$

$$\lambda(t) = \lambda_*(t), \qquad \rho(t) = \rho_*(t), \qquad \mu = \mu_* . \tag{107}$$

The implication of (106)-(107) is the following: under the assumption of linear constraints, the general solution of (70)-(76) can now be obtained by executing only one sweep

(instead of two) of $n+p+1$ integrations, namely, the sweep necessary to generate the particular solution (94). This is the solution corresponding to (93-1), namely, the solution associated with the ordinary gradient method of Ref. 9.

Isoperimetric Constant. Under the linearity assumption for the constraints, Eq.(79) still holds, but the error in the optimality conditions (78) simplifies to

$$Q = \int_0^1 B_\star^T B_\star dt + C_\star^T C_\star \ . \tag{108}$$

Descent Property. Under the linearity assumption for the constraints, Eq.(83) still holds, but the functional (82) simplifies to

$$Z = \int_0^1 B_\star^T \hat{B} dt + C_\star^T \hat{C} \ . \tag{109}$$

Local Orthogonality Conditions. Under the linearity assumption for the constraints, the two-parameter family (99) simplifies to

$$\tilde{x}(t) = x(t) + \alpha[A_\star(t) + \gamma\hat{A}(t)], \tag{110-1}$$

$$\tilde{u}(t) = u(t) + \alpha[B_\star(t) + \gamma\hat{B}(t)], \tag{110-2}$$

$$\tilde{\pi} = \pi + \alpha(C_\star + \gamma\hat{C}) \ . \tag{110-3}$$

Next, we consider the augmented functional (98) and observe that, for the two-parameter family (110), it takes the form (100). Therefore, the optimal values of the stepsize α and the directional coefficient γ satisfy the relations (101). Because of the assumed linearity of the constraints (13)-(15), Eqs. (101) take the particular form

$$\int_0^1 (\tilde{f}_x - \phi_x\lambda + S_x\rho - \dot{\lambda})^T A\,dt + \int_0^1 (\tilde{f}_u - \phi_u\lambda + S_u\rho)^T B\,dt$$
$$+\left[\int_0^1 (\tilde{f}_\pi - \phi_\pi\lambda + S_\pi\rho)\,dt + (\tilde{g}_\pi + \psi_\pi\mu)_1\right]^T C + \left[(\lambda + \tilde{g}_x + \psi_x\mu)^T A\right]_1 = 0, \tag{111}$$

$$\int_0^1 (\tilde{f}_x - \phi_x\lambda + S_x\rho - \dot{\lambda})^T \hat{A}\,dt + \int_0^1 (\tilde{f}_u - \phi_u\lambda + S_u\rho)^T \hat{B}\,dt$$
$$+\left[\int_0^1 (\tilde{f}_\pi - \phi_\pi\lambda + S_\pi\rho)\,dt + (\tilde{g}_\pi + \psi_\pi\mu)_1\right]^T \hat{C} + \left[(\lambda + \tilde{g}_x + \psi_x\mu)^T \hat{A}\right]_1 = 0, \tag{112}$$

with the implication that

$$\int_0^1 (\tilde{f}_x - \phi_x\lambda + S_x\rho - \dot{\lambda})^T A_\star\,dt + \int_0^1 (\tilde{f}_u - \phi_u\lambda + S_u\rho)^T B_\star\,dt$$
$$+\left[\int_0^1 (\tilde{f}_\pi - \phi_\pi\lambda + S_\pi\rho)\,dt + (\tilde{g}_\pi + \psi_\pi\mu)_1\right]^T C_\star + \left[(\lambda + \tilde{g}_x + \psi_x\mu)^T A_\star\right]_1 = 0. \tag{113}$$

Because of the linearity of the constraints (13)-(15) and after invoking Eqs. (70)-(72), one can show that Eqs. (111)-(113) hold for any distribution of Lagrange multipliers. In particu-

lar, they hold if $\lambda(t)$, $\rho(t)$, μ are replaced by $\tilde{\lambda}(t)$, $\tilde{\rho}(t)$, $\tilde{\mu}$. This yields the supplementary relations

$$\int_0^1 (\tilde{f}_x - \phi_x\tilde{\lambda} + S_x\tilde{\rho} - \dot{\tilde{\lambda}})^T A dt + \int_0^1 (\tilde{f}_u - \phi_u\tilde{\lambda} + S_u\tilde{\rho})^T B dt$$
$$+ \left[\int_0^1 (\tilde{f}_\pi - \phi_\pi\tilde{\lambda} + S_\pi\tilde{\rho}) dt + (\tilde{g}_\pi + \psi_\pi\tilde{\mu})_1\right]^T C + \left[(\tilde{\lambda} + \tilde{g}_x + \psi_x\tilde{\mu})^T A\right]_1 = 0, \quad (114)$$

$$\int_0^1 (\tilde{f}_x - \phi_x\tilde{\lambda} + S_x\tilde{\rho} - \dot{\tilde{\lambda}})^T \hat{A} dt + \int_0^1 (\tilde{f}_u - \phi_u\tilde{\lambda} + S_u\tilde{\rho})^T \hat{B} dt$$
$$+ \left[\int_0^1 (\tilde{f}_\pi - \phi_\pi\tilde{\lambda} + S_\pi\tilde{\rho}) dt + (\tilde{g}_\pi + \psi_\pi\tilde{\mu})_1\right]^T \hat{C} + \left[(\tilde{\lambda} + \tilde{g}_x + \psi_x\tilde{\mu})^T \hat{A}\right]_1 = 0, \quad (115)$$

$$\int_0^1 (\tilde{f}_x - \phi_x\tilde{\lambda} + S_x\tilde{\rho} - \dot{\tilde{\lambda}})^T A_* dt + \int_0^1 (\tilde{f}_u - \phi_u\tilde{\lambda} + S_u\tilde{\rho})^T B_* dt$$
$$+ \left[\int_0^1 (\tilde{f}_\pi - \phi_\pi\tilde{\lambda} + S_\pi\tilde{\rho}) dt + (\tilde{g}_\pi + \psi_\pi\tilde{\mu})_1\right]^T C_* + \left[(\tilde{\lambda} + \tilde{g}_x + \psi_x\tilde{\mu})^T A_*\right]_1 = 0. \quad (116)$$

Next, we combine Eqs. (114)-(116) with Eqs. (73)-(76) written for the next iteration. This leads to the following local orthogonality conditions:

$$\int_0^1 \tilde{B}_*^T B dt + \tilde{C}_*^T C = 0, \quad (117\text{-}1)$$

$$\int_0^1 \tilde{B}_*^T \hat{B} dt + \tilde{C}_*^T \hat{C} = 0, \quad (117\text{-}2)$$

$$\int_0^1 \tilde{B}_*^T B_* dt + \tilde{C}_*^T C_* = 0. \quad (117\text{-}3)$$

Here, the adjective local is employed to mean that Eqs. (117) involve vectors $B(t)$, C which are solutions of (70)-(76) computed for the present iteration and the previous iteration; they also involve vectors $B_*(t)$, C_* which are solutions of (70)-(76) for $\gamma = 0$ computed for the present iteration and the next iteration.

7. Conjugate Gradient Phase: Quadratic Functional and Linear Constraints

In the previous section, we assumed that the constraints (13)-(15) are linear and arrived at the local orthogonality conditions (117). In this section, we retain the constraint linearity hypothesis and further asssume that the functional (12) is quadratic.

For the sake of compactness, let y and E denote the vectors

$$y = \begin{bmatrix} x \\ u \\ \pi \end{bmatrix}, \qquad E = \begin{bmatrix} A \\ B \\ C \end{bmatrix}. \tag{118}$$

Let f_y and g_y denote the gradients of the functions f and g with respect to the vector y:

$$f_y = \begin{bmatrix} f_x \\ f_u \\ f_\pi \end{bmatrix}, \qquad g_y = \begin{bmatrix} g_x \\ 0 \\ g_\pi \end{bmatrix}. \tag{119}$$

Under the assumption that the functions f and g are quadratic in their respective arguments, the following exact relations can be established:

$$\tilde{f}_y = f_y + \alpha f_{yy} E, \qquad \tilde{g}_y = g_y + \alpha g_{yy} E, \tag{120}$$

where

$$f_{yy} = \begin{bmatrix} f_{xx} & f_{xu} & f_{x\pi} \\ f_{ux} & f_{uu} & f_{u\pi} \\ f_{\pi x} & f_{\pi u} & f_{\pi\pi} \end{bmatrix}, \qquad g_{yy} = \begin{bmatrix} g_{xx} & 0 & g_{x\pi} \\ 0 & 0 & 0 \\ g_{\pi x} & 0 & g_{\pi\pi} \end{bmatrix}. \tag{121}$$

With this understanding, Eqs.(111)-(113) become

$$\begin{aligned} &\int_0^1 (f_x - \phi_x\lambda + S_x\rho - \dot{\lambda})^T A dt + \int_0^1 (f_u - \phi_u\lambda + S_u\rho)^T B dt \\ &+ \left[\int_0^1 (f_\pi - \phi_\pi\lambda + S_\pi\rho) dt + (g_\pi + \psi_\pi\mu)_1\right]^T C + \left[(\lambda + g_x + \psi_x\mu)^T A\right]_1 \\ &+\alpha\left[\int_0^1 E^T f_{yy} E dt + (E^T g_{yy} E)_1\right] = 0, \end{aligned} \tag{122}$$

$$\begin{aligned} &\int_0^1 (f_x - \phi_x\lambda + S_x\rho - \dot{\lambda})^T \hat{A} dt + \int_0^1 (f_u - \phi_u\lambda + S_u\rho)^T \hat{B} dt \\ &+ \left[\int_0^1 (f_\pi - \phi_\pi\lambda + S_\pi\rho) dt + (g_\pi + \psi_\pi\mu)_1\right]^T \hat{C} + \left[(\lambda + g_x + \psi_x\mu)^T \hat{A}\right]_1 \\ &+\alpha\left[\int_0^1 E^T f_{yy} \hat{E} dt + (E^T g_{yy} \hat{E})_1\right] = 0, \end{aligned} \tag{123}$$

$$\begin{aligned} &\int_0^1 (f_x - \phi_x\lambda + S_x\rho - \dot{\lambda})^T A_* dt + \int_0^1 (f_u - \phi_u\lambda + S_u\rho)^T B_* dt \\ &+ \left[\int_0^1 (f_\pi - \phi_\pi\lambda + S_\pi\rho) dt + (g_\pi + \psi_\pi\mu)_1\right]^T C_* + \left[(\lambda + g_x + \psi_x\mu)^T A_*\right]_1 \\ &+\alpha\left[\int_0^1 E^T f_{yy} E_* dt + (E^T g_{yy} E_*)_1\right] = 0. \end{aligned} \tag{124}$$

Upon invoking Eqs. (73)-(76), we see that Eqs. (122)-(124) can be rewritten as

$$\int_0^1 B_\star^T B dt + C_\star^T C - \alpha \left[\int_0^1 E^T f_{yy} E dt + (E^T g_{yy} E)_1 \right] = 0, \tag{125-1}$$

$$\int_0^1 B_\star^T \hat{B} dt + C_\star^T \hat{C} - \alpha \left[\int_0^1 E^T f_{yy} \hat{E} dt + (E^T g_{yy} \hat{E})_1 \right] = 0, \tag{125-2}$$

$$\int_0^1 B_\star^T B_\star dt + C_\star^T C_\star - \alpha \left[\int_0^1 E^T f_{yy} E_\star dt + (E^T g_{yy} E_\star)_1 \right] = 0. \tag{125-3}$$

Local Conjugacy Condition. Next, we employ the local orthogonality condition (117-1) written for the present iteration, and observe that (125-2) yields the local conjugacy condition

$$\int_0^1 E^T f_{yy} \hat{E} dt + (E^T g_{yy} \hat{E})_1 = 0. \tag{126}$$

Here, the adjective local is employed to mean that Eq. (126) involves vectors $E(t)$, that is, vectors $A(t)$, $B(t)$, C, which are solutions of (70)-(76) computed for the present iteration and the previous iteration.

Stepsize. After observing that

$$E = E_\star + \gamma \hat{E} \tag{127}$$

and making use of the local conjugacy condition (126), we see

that Eq.(125-3) can be rewritten as

$$\int_0^1 B_\star^T B_\star dt + C_\star^T C_\star - \alpha \left[\int_0^1 E^T f_{yy} E dt + (E^T g_{yy} E)_1\right] = 0. \tag{128}$$

This equation enables one to compute the optimum stepsize α, once the value of the directional coefficient γ is known.

Directional Coefficient. After invoking Eq. (127), we see that the local conjugacy condition (126) becomes

$$\int_0^1 E_\star^T f_{yy} \hat{E} dt + (E_\star^T g_{yy} \hat{E})_1 + \gamma \left[\int_0^1 \hat{E}^T f_{yy} \hat{E} dt + (\hat{E}^T g_{yy} \hat{E})_1\right] = 0. \tag{129}$$

If we employ Eq. (128) written for the previous iteration, Eq. (129) becomes

$$\gamma \left[\int_0^1 \hat{B}_\star^T \hat{B}_\star dt + \hat{C}_\star^T \hat{C}_\star\right] + \hat{\alpha} \left[\int_0^1 E_\star^T f_{yy} \hat{E} dt + (E_\star^T g_{yy} \hat{E})_1\right] = 0 \tag{130}$$

and, in the light of Eqs. (120), can be rewritten as

$$\gamma \left[\int_0^1 \hat{B}_\star^T \hat{B}_\star dt + \hat{C}_\star^T \hat{C}_\star\right] + \int_0^1 E_\star^T (f_y - \hat{f}_y) dt + \left[E_\star^T (g_y - \hat{g}_y)\right]_1 = 0. \tag{131}$$

If Eqs. (70)-(76) and (117)-(119) are employed, the following relations can be shown to hold:

$$\int_0^1 E_\star^T f_y dt + (E_\star^T g_y)_1 + \int_0^1 B_\star^T B_\star dt + C_\star^T C_\star = 0 , \tag{132}$$

$$\int_0^1 E_*^T \hat{f}_y \, dt + (E_*^T \hat{g}_y)_1 = 0 \ . \tag{133}$$

As a consequence, Eq.(131) becomes

$$\gamma \left[\int_0^1 \hat{B}_*^T \hat{B}_* dt + \hat{C}_*^T \hat{C}_* \right] - \left[\int_0^1 B_*^T B_* dt + C_*^T C_* \right] = 0 \tag{134}$$

and can be rewritten as

$$\gamma = Q/\hat{Q} \ , \tag{135}$$

where

$$Q = \int_0^1 B_*^T B_* dt + C_*^T C_* \ , \tag{136}$$

$$\hat{Q} = \int_0^1 \hat{B}_*^T \hat{B}_* dt + \hat{C}_*^T \hat{C}_* \tag{137}$$

denote the errors in the optimality conditions for the present iteration and the previous iteration, respectively. These quantities are known, since they involve vectors $B_*(t)$, C_* which are solutions of (70)-(76) for $\gamma = 0$ computed for the present iteration and the previous iteration.

Descent Property. Because of the local orthogonality condition (117-1) written for the previous iteration, Eq. (109) yields

$$Z = 0. \tag{138}$$

As a consequence, the first variation of the augmented functional (83) reduces to

$$\delta J = -\alpha Q \ , \tag{139}$$

where the error in the optimality condition Q is given by Eq. (136). Equation (139) holds for any conjugate gradient iteration and shows that, since $Q > 0$, we have $\delta J < 0$. Hence, for α sufficiently small, a decrease in the augmented functional J is guaranteed. In conclusion, for the linear-quadratic case, the restart procedure mentioned in Section 5 never occurs. This means that the directional coefficient γ is set at the level (84) only for the first iteration.

General Orthogonality and Conjugacy Conditions. Assume now that the algorithm described by Eqs. (70)-(76) and (110) is employed, starting with a feasible nominal solution. Further, assume that the first conjugate gradient iteration is done with

$$\gamma = 0, \tag{140}$$

meaning that this is an ordinary gradient iteration. Under these assumptions and with reference to the linear-quadratic case, one can generalize the local orthogonality conditions (117) and the local conjugacy condition (126) as follows:

$$\int_0^1 B_*^T B_p dt + C_*^T C_p = 0 \ , \tag{141-1}$$

$$\int_0^1 B_*^T B_{*p} dt + C_*^T C_{*p} = 0 \ , \tag{141-2}$$

and

$$\int_0^1 E^T f_{yy} E_p dt + (E^T g_{yy} E_p)_1 = 0 , \tag{142}$$

where the subscript p denotes any iteration preceding the present iteration. While these equations do not guarantee convergence in a finite number of steps, they do guarantee that the algorithm generates a sequence of linearly indepentent vectors $E(t)$, that is, a sequence of linearly independent variations per unit stepsize $A(t)$, $B(t)$, C.

8. Conjugate Gradient Phase: Practical Implementation

In this section, we summarize the results of Sections 5-7, and suggest practical ways of utilizing these results, while avoiding the use of second derivatives. This is an essential characteristic of a first-order method.

Auxiliary Functions. The first step is to solve Eqs. (70)-(76) for a fictitious value of the directional coefficient, namely,

$$\gamma_* = 0 \ . \tag{143}$$

This yields the following linear, two-point boundary-value problem:

$$\dot{A}_* - \phi_x^T A_* - \phi_u^T B_* - \phi_\pi^T C_* = 0, \qquad 0 \le t \le 1, \tag{144}$$

$$S_x^T A_* + S_u^T B_* + S_\pi^T C_* = 0, \qquad 0 \le t \le 1, \tag{145}$$

$$(A_*)_0 = 0 \ , \qquad (\psi_x^T A_* + \psi_\pi^T C_*)_1 = 0, \tag{146}$$

$$\dot{\lambda}_* - f_x + \phi_x \lambda_* - S_x \rho_* = 0, \qquad 0 \le t \le 1, \tag{147}$$

$$B_* + f_u - \phi_u \lambda_* + S_u \rho_* = 0, \qquad 0 \le t \le 1, \tag{148}$$

$$C_* + \int_0^1 (f_\pi - \phi_\pi \lambda_* + S_\pi \rho_*)dt + (g_\pi + \psi_\pi \mu_*)_1 = 0, \tag{149}$$

$$(\lambda_* + g_x + \psi_x \mu_*)_1 = 0. \tag{150}$$

Using the solution technique of Section 5, we obtain the following auxiliary functions and multipliers:[8]

$$A_*(t),\ B_*(t),\ C_*,\lambda_*(t),\ \rho_*(t),\ \mu_* \quad . \tag{151}$$

Directional Coefficient. The second step is to compute the actual value of the directional coefficient γ. For the first conjugate gradient phase, we set

$$\gamma = 0, \tag{152}$$

meaning that the conjugate gradient iteration is an ordinary gradient iteration. For subsequent conjugate gradient phases, we set

$$\gamma = Q/\hat{Q} \ , \tag{153}$$

where

$$Q = \int_0^1 B_*^T B_* dt + C_*^T C_* \ , \tag{154-1}$$

$$\hat{Q} = \int_0^1 \hat{B}_*^T \hat{B}_* dt + \hat{C}_*^T \hat{C}_* \tag{154-2}$$

denote the errors in the optimality conditions at the beginning

[8] These functions and multipliers are identical with those solving the linear, two-point boundary-value problem associated with the ordinary gradient phase of Ref. 9.

of the present conjugate gradient phase and at the beginning of the previous conjugate gradient phase, respectively.

Note that the directional coefficient (153) is acceptable only if

$$\tilde{J}_\alpha(0) < 0 , \tag{155}$$

where $\tilde{J}_\alpha(0)$ is given by Eq. (160). If Ineq. (155) is violated, then the directional coefficient (153) must be discarded and replaced by the value (152). This means that the algorithm must be restarted by replacing the conjugate gradient phase with an ordinary gradient phase.

<u>Basic Functions</u>. The third step is to compute the basic functions A(t), B(t), C and the multipliers $\lambda(t)$, $\rho(t)$, μ. This is done with the following formulas:

$$A(t) = A_\star(t) + \gamma\hat{A}(t), \qquad B(t) = B_\star(t) + \gamma\hat{B}(t), \qquad C = C_\star + \gamma\hat{C}, \tag{156}$$

$$\lambda(t) = \lambda_\star(t), \qquad \rho(t) = \rho_\star(t), \qquad \mu = \mu_\star . \tag{157}$$

<u>Stepsize</u>. With the basic functions A(t), B(t), C known, we consider the one-parameter family of solutions (30). For this one-parameter family, the augmented functional (16) takes the form

$$\tilde{J}(\alpha) = (\lambda^T\tilde{x})_0^1 + \int_0^1 (\tilde{f} - \lambda^T\tilde{\phi} + \rho^T\tilde{S} - \dot{\lambda}^T\tilde{x})dt + (\tilde{g} + \mu^T\tilde{\psi})_1 , \tag{158}$$

with the implication that

$$\tilde{J}_\alpha(\alpha) = \int_0^1 (\tilde{f}_x - \tilde{\phi}_x\lambda + \tilde{S}_x\rho - \dot{\lambda})^T A dt + \int_0^1 (\tilde{f}_u - \tilde{\phi}_u\lambda + \tilde{S}_u\rho)^T B dt$$

$$+ \left[\int_0^1 (\tilde{f}_\pi - \tilde{\phi}_\pi\lambda + \tilde{S}_\pi\rho)dt + (\tilde{g}_\pi + \tilde{\psi}_\pi\mu)_1\right]^T C + \left[(\lambda + \tilde{g}_x + \tilde{\psi}_x\mu)^T A\right]_1 , \quad (159)$$

and with the further implication that

$$\tilde{J}_\alpha(0) = -(Q + \gamma Z) , \quad (160)$$

where

$$Q = \int_0^1 B_*^T B_* dt + C_*^T C_* , \quad (161)$$

$$Z = \int_0^1 B_*^T \hat{B} dt + C_*^T \hat{C} . \quad (162)$$

Note that

$$Z = 0 \quad (163)$$

in the linear-quadratic case and that

$$Z \neq 0 \quad (164)$$

in the general case.

In addition to the augmented functional, the constraint error (24) must be monitored during the conjugate gradient phase. For the one-parameter family of solutions (30), the

functional (24) takes the form

$$\tilde{P}(\alpha) = \int_0^1 N(\dot{x} - \tilde{\phi})dt + \int_0^1 N(\tilde{S})dt + N(\tilde{\psi})_1 \ , \tag{165}$$

where N(b) denotes the norm squared of the vector b.

With these preliminaries in mind, the stepsize α must be selected so that the following inequalities are satisfied:

$$\tilde{J}_\alpha^2(\alpha)/\tilde{J}_\alpha^2(0) \le \varepsilon_3 \ , \tag{166}$$

$$\tilde{J}(\alpha) < \tilde{J}(0) \ , \tag{167}$$

subject to

$$\tilde{P}(\alpha) \le P_* \ , \qquad \tilde{\tau}(\alpha) \ge 0 \ , \tag{168}$$

where ε_3 and P_* are preselected numbers. While ε_3 is a small number, P_* need not be necessarily small.

Quadratic Interpolation. In the linear-quadratic case, satisfaction of (166)-(167) can be achieved by using quadratic interpolation. The procedure is as follows.

Let the function (158) be written in the quadratic form

$$\tilde{J}(\alpha) = k_0 + k_1\alpha + k_2\alpha^2 \ , \tag{169}$$

with the implication that

$$\tilde{J}_\alpha(\alpha) = k_1 + 2k_2\alpha \ , \qquad \tilde{J}_{\alpha\alpha}(\alpha) = 2k_2 \ . \tag{170}$$

Therefore, the coefficients k_0 and k_1 are given by

$$k_0 = \tilde{J}(0), \qquad k_1 = \tilde{J}_\alpha(0) . \tag{171}$$

The coefficient k_2 can be computed by evaluating the augmented functional at some reference stepsize, for instance, $\alpha = 1$. If this is done, one concludes that

$$k_2 = \tilde{J}(1) - \tilde{J}(0) - \tilde{J}_\alpha(0) . \tag{172}$$

With the coefficients known, the optimum value of α can be computed from the relation

$$\tilde{J}_\alpha(\alpha) = 0 , \tag{173}$$

which implies that

$$\alpha_0 = - k_1/2k_2 . \tag{174}$$

In the linear quadratic case, the representation of the augmented functional (158) by means of the quadratic form (169) is exact. Therefore, the quadratic interpolation process is employed only once (one-step quadratic interpolation).

Cubic Interpolation. In the general case, it is better to try to achieve satisfaction of (166)-(167) with cubic interpolation. The procedure is as follows.

We consider the reference stepsize k and the sequence of stepsizes

$$\{\alpha\} = \{0,\ k,\ 2k,\ 4k,\ 8k, \ldots\} \quad . \tag{175}$$

For every element of the sequence (175), we compute the quantities (158)-(159). We denote by α_1 and α_2 the smallest consecutive elements in the sequence (175) such that the following inequalities are satisfied

$$\tilde{J}_\alpha(\alpha_1) < 0 \ , \quad \tilde{J}_\alpha(\alpha_2) > 0 \ . \tag{176}$$

Then, assuming that the derivative $\tilde{J}_\alpha(\alpha)$ is continuous, a relative minimum of $\tilde{J}(\alpha)$ occurs for a value α_0 such that

$$\alpha_1 < \alpha_0 < \alpha_2 \ . \tag{177}$$

In order to find the minimum of $\tilde{J}(\alpha)$ numerically, we approximate the function $\tilde{J}(\alpha)$ with the cubic form

$$\tilde{J}(\alpha) = k_0 + k_1\alpha + k_2\alpha^2 + k_3\alpha^3 \ , \tag{178}$$

with the implication that

$$\tilde{J}_\alpha(\alpha) = k_1 + 2k_2\alpha + 3k_3\alpha^2 \ , \qquad \tilde{J}_{\alpha\alpha}(\alpha) = 2k_2 + 6k_3\alpha \ . \tag{179}$$

The coefficients k_i are computed by forcing the cubic function (178) and its derivative to satisfy the exact values of the ordinate $\tilde{J}(\alpha)$ and the slope $\tilde{J}_\alpha(\alpha)$ at α_1 and α_2; that is, the coefficients k_i are computed from the conditions

$$\tilde{J}(\alpha_1) = k_0 + k_1\alpha_1 + k_2\alpha_1^2 + k_3\alpha_1^3 \ , \tag{180-1}$$

$$\tilde{J}(\alpha_2) = k_0 + k_1\alpha_2 + k_2\alpha_2^2 + k_3\alpha_2^3 \ , \tag{180-2}$$

$$\tilde{J}_\alpha(\alpha_1) = k_1 + 2k_2\alpha_1 + 3k_3\alpha_1^2 \ , \tag{180-3}$$

$$\tilde{J}_\alpha(\alpha_2) = k_1 + 2k_2\alpha_2 + 3k_3\alpha_2^2 \ . \tag{180-4}$$

With the coefficients known, the optimum value of α can be computed from the relation

$$\tilde{J}_\alpha(\alpha) = 0 \ , \tag{181}$$

which implies that

$$\alpha_0 = (1/3k_3)\left[-k_2 + \surd(k_2^2 - 3k_1k_3)\right] . \tag{182}$$

Then, two possibilities arise, depending on the sign of $\tilde{J}_\alpha(\alpha)$ at the point α_0:

$$\text{(i)} \qquad \tilde{J}_\alpha(\alpha_0) > 0 \ , \tag{183-1}$$

$$\text{(ii)} \qquad \tilde{J}_\alpha(\alpha_0) < 0 \ . \tag{183-2}$$

In Case (i), the cubic interpolation process is repeated between α_1 and α_0. In Case (ii), it is repeated between α_0 and α_2. The process is continued until Ineq. (166) is satisfied.

Limiting Case. If the solution of Eqs. (180) is such that

$$k_3 = 0 , \tag{184}$$

the optimal stepsize α_0 cannot be computed with Eq. (182), since both the numerator and the denominator vanish simultaneously. This difficulty can be bypassed by observing that the limiting case (184) means that the cubic approximation (178) is being replaced by the quadratic approximation (169). As a consequence, the optimal stepsize of the cubic approximation (182) must be replaced by the optimal stepsize of the quadratic approximation (174).

In practice, two cases are possible:

$$\text{(i)} \quad k_3^2 > \varepsilon_4 , \tag{185-1}$$

$$\text{(ii)} \quad k_3^2 \leq \varepsilon_4 , \tag{185-2}$$

where ε_4 is a small, preselected number. In Case (i), the optimal stepsize α_0 must be computed with Eq. (182). In Case (ii), the optimal stepsize α_0 must be computed with (174).

Remark. For more details concerning the one-dimensional search for the conjugate gradient stepsize, the reader is referred to Ref. 16.

9. Descent Property of a Cycle

While the stepsizes employed in the conjugate gradient phase and the restoration phase are not necessarily small, a descent property can be proven for a complete conjugate gradient-restoration cycle under the assumption of small stepsizes.

Let the subscript g denote the conjugate gradient phase, and let the subscript r denote the restoration phase. Simple manipulations, omitted for the sake of brevity, show that the conjugate gradient corrections and the restoration corrections have the following orders of magnitude:

$$\Delta x_g(t) = O(\alpha_g), \qquad \Delta u_g(t) = O(\alpha_g), \qquad \Delta\pi_g = O(\alpha_g), \tag{186}$$

$$\Delta x_r(t) = O(\alpha_r \alpha_g^2), \qquad \Delta u_r(t) = O(\alpha_r \alpha_g^2), \qquad \Delta\pi_r = O(\alpha_r \alpha_g^2). \tag{187}$$

For α_g sufficiently small, the restoration corrections are negligible with respect to the conjugate gradient corrections. Therefore, providing the conjugate gradient phase has a descent property on $\tilde{J}(\alpha)$ (this is guaranteed through the selection of the directional coefficient γ), the restoration phase preserves the descent property of the conjugate gradient phase.

More specifically, let the subscripts 1, 2, 3 denote the values of the functional I at the beginning of the conjugate gradient phase, at the end of the conju-

gate gradient phase, and at the end of the subsequent restoration phase. We note that I_1 and I_2 are not comparable, since the constraints are not satisfied to the same accuracy. On the other hand, not only I_1 and I_3 are comparable, but the conjugate gradient stepsize α_g can be selected so that

$$I_3 < I_1 \ . \tag{188}$$

This constitutes the descent property of a complete conjugate gradient-restoration cycle.

If Ineq. (188) is satisfied, the next conjugate gradient-restoration cycle can be started. If Ineq. (188) is violated, one must return to the previous conjugate gradient phase and bisect the conjugate gradient stepsize α_g until, after restoration, Ineq. (188) is satisfied. That the above procedure leads to satisfaction of Ineq. (188) is guaranteed by two circumstances: first, the fact that the directional coefficient γ has been chosen consistently with Ineq. (155); second, the fact that, for α_g small, the restoration corrections (187) are negligible by comparison with the conjugate gradient corrections (186).

10. Summary of the Algorithm

A sequential conjugate gradient-restoration algorithm has been developed in order to solve optimal control problems involving a functional (12), subject to differential constraints (13), nondifferential constraints (14), and terminal constraints(15). The algorithm is composed of a sequence of cycles, each cycle consisting of two phases, a conjugate gradient phase and a restoration phase.

Decision Variables. The major decision variables controlling the algorithm are the constraint error P, given by Eq. (24), and the optimality condition error Q, given by Eq. (25).

Depending on the value of P, two cases are possible:

$$\text{(i)} \qquad P > \varepsilon_1 , \tag{189}$$

$$\text{(ii)} \qquad P \leq \varepsilon_1 . \tag{190}$$

In Case (i), the algorithm executes a restoration phase. In Case (ii), the algorithm computes the optimality condition error Q.

Depending on the value of Q, two subcases of Case (ii) are possible:

$$\text{(iii)} \qquad P \leq \varepsilon_1 , \qquad Q > \varepsilon_2 , \tag{191}$$

$$\text{(iv)} \qquad P \leq \varepsilon_1 , \qquad Q \leq \varepsilon_2 . \tag{192}$$

In Case (iii), the algorithm executes the conjugate gradient phase. In Case (iv), the algorithm stops: convergence has been achieved.

Iterations. Each iteration of the conjugate gradient phase or the restoration phase is described by the following relations:

$$\tilde{x}(t) = x(t) + \alpha A(t) , \qquad \tilde{u}(t) = u(t) + \alpha B(t) , \qquad \tilde{\pi} = \pi + \alpha C, \tag{193}$$

which tie the nominal functions and the varied functions. Therefore, each iteration includes two distinct operations: the determination of the basic functions $A(t)$, $B(t)$, C and the determination of the stepsize α .

Restoration Phase. The restoration phase includes one or more restorative iterations. A restorative iteration is started whenever the constraint error P satisfies Ineq. (189).

In each restorative iteration, the basic functions $A(t)$, $B(t)$, C are determined by solving the linear, two-point boundary-value problem (43)-(49) using the method of particular solutions. This requires executing $n + p + 1$ independent sweeps of the system (43)-(49).

The stepsize α must be determined so that the following

inequalities are satisfied:

$$\tilde{P}(\alpha) < \tilde{P}(0), \qquad \tilde{\tau}(\alpha) \geq 0 \ . \tag{194}$$

For this purpose, a bisection process, starting from the reference stepsize

$$\alpha_0 = 1 , \tag{195}$$

is employed.

In the course of a restorative iteration, the reduction of the constraint error is guaranteed. However, there is no guarantee that the constraint error is reduced below the threshold (190) characterizing the beginning of the next conjugate gradient phase. In other words, after Ineqs. (194) have been satisfied, two cases are possible:

$$\text{(i)} \qquad \tilde{P}(\alpha) > \varepsilon_1 , \tag{196}$$

$$\text{(ii)} \qquad \tilde{P}(\alpha) \leq \varepsilon_1 . \tag{197}$$

In Case (i), a further restorative iteration is initiated employing as nominal functions the varied functions of the previous restorative iteration. In Case (ii), the restoration phase is terminated, and the next conjugate gradient phase is started.

Clearly, each restoration phase includes a variable

number of restorative iterations, depending on the particular problem and the nominal functions employed. Generally speaking, the number of restorative iterations per restoration phase decreases in subsequent cycles of the sequential conjugate gradient-restoration algorithm and approaches zero as the algorithm proceeds toward convergence.

<u>Conjugate Gradient Phase</u>. The conjugate gradient phase involves a single iteration. This single iteration is started whenever the constraint error P satisfies Ineq. (190).

In each conjugate gradient iteration, the first step is to compute the auxiliary functions $A_*(t)$, $B_*(t)$, C_* corresponding to the fictitious value

$$\gamma_* = 0 \tag{198}$$

of the directional coefficient. These auxiliary functions are determined by solving the linear, two-point boundary-value problem (144)-(150) using the method of particular solutions. Once more, this requires executing $n + p + 1$ independent sweeps of the system (144)-(150).

With the auxiliary functions known, the basic functions are determined with the relations

$$A(t) = A_*(t) + \gamma\hat{A}(t), \qquad B(t) = B_*(t) + \gamma\hat{B}(t), \qquad C = C_* + \gamma\hat{C}, \tag{199}$$

where γ denotes the actual value of the directional coefficient. This directional coefficient is set at one of

the following levels:

$$\text{(i)} \qquad \gamma = 0\,, \tag{200}$$

$$\text{(ii)} \qquad \gamma = Q/\hat{Q}\,, \tag{201}$$

where (200) holds for the first conjugate gradient iteration and (201) holds for any subsequent conjugate gradient iteration.

Prior to accepting the directional coefficient (200) or (201), a check must be made. For the choice (200) or (201), is the slope of the augmented functional $\tilde{J}_\alpha(0)$ negative? In other words, does the descent property of the gradient phase hold?

Concerning Case (i), two subcases are possible:

$$\text{(iii)} \qquad \gamma = 0\,, \qquad \tilde{J}_\alpha(0) < 0\,, \tag{202}$$

$$\text{(iv)} \qquad \gamma = 0\,, \qquad \tilde{J}_\alpha(0) \geq 0\,. \tag{203}$$

In Case (iii), the directional coefficient (200) is accepted, and the algorithm completes the ordinary gradient phase. In Case (iv), the descent property of the ordinary gradient phase does not hold, and the value of the augmented functional cannot be reduced, owing to numerical inaccuracies; hence, this constitutes a nonconvergence condition for the algorithm as a whole.

Concerning Case (ii), two subcases are possible:

$$\text{(v)} \qquad \gamma = Q/\hat{Q}\,, \qquad \tilde{J}_\alpha(0) < 0\,, \tag{204}$$

$$\text{(vi)} \qquad \gamma = Q/\hat{Q}\,, \qquad \tilde{J}_\alpha(0) \geq 0\,. \tag{205}$$

In Case (v), the directional coefficient (201) is accepted, and the algorithm completes the conjugate gradient phase. In Case (vi), the directional coefficient (201) is rejected, and is replaced by the directional coefficient (200). This means that the algorithm is restarted with an ordinary gradient phase, characterized by $\gamma = 0$.

After the directional coefficient has been selected, the functions (199) are known, and the one-parameter family of solutions (193) can be formed. Then, the conjugate gradient stepsize α must be determined through a one-dimensional search on the augmented functional $\tilde{J}(\alpha)$ in such a way that the following inequalities are satisfied:[9,10]

$$\tilde{J}^2_\alpha(\alpha)/\tilde{J}^2_\alpha(0) \leq \varepsilon_3 \tag{206}$$

---

[9] Note that Ineq. (208) prevents the constraint error P from becoming too large during the conjugate gradient phase.

[10] For the details of the one-dimensional search technique leading to the satisfaction of (207)-(209), see Ref. 16.

and

$$\tilde{J}(\alpha) < \tilde{J}(0), \tag{207}$$

$$\tilde{P}(\alpha) \leq P_* , \tag{208}$$

$$\tilde{\tau}(\alpha) \geq 0 . \tag{209}$$

For this purpose, the cubic interpolation procedure of Section 8, is employed, with an optional switch to quadratic interpolation, if needed.

<u>Conjugate Gradient-Restoration Cycle</u>. Generally speaking, the first cycle of the algorithm is a half cycle, in that it includes a restoration phase only. Every subsequent cycle is a complete cycle, in that it includes both a conjugate gradient phase and a restoration phase.

Between the endpoints of a complete conjugate gradient-restoration cycle, the following descent property must be satisfied:

$$I_3 < I_1 , \tag{210}$$

where I_1 denotes the value of the functional (12) at the beginning of the cycle and I_3 denotes the value of (12) at the end of the cycle.

If Ineq. (210) holds, the next conjugate gradient-restoration cycle can be started. If Ineq. (210) is violated,

one must return to the previous conjugate gradient phase and bisect the conjugate gradient stepsize until, after restoration, Ineq. (210) is satisfied.

11. Remarks and Safeguards

In this final section, we include miscellaneous considerations, relevant to the computer implementation of the sequential conjugate gradient-restoration algorithm. We also list some important safeguards.

Nondifferential Constraint. For the restoration phase, the linear, two-point boundary-value problem (43)-(49) must be solved. During the execution of a sweep, assume that time station t has been reached and that $A(t)$, C, $\lambda(t)$ are known at that time station. Then, Eqs. (44) and (47) constitute a system of $m+k$ equations in the $m+k$ components of the vectors $B(t)$, $\rho(t)$. The system admits a unique solution providing the following relation is satisfied:

$$\det \begin{bmatrix} I & S_u \\ S_u^T & 0 \end{bmatrix} = (-1)^k \det \begin{bmatrix} S_u^T S_u \end{bmatrix} \neq 0 , \tag{211}$$

where I denotes the m x m identity matrix and 0 denotes the k x k null matrix.

An analogous remak holds for the conjugate gradient phase: here, the linear, two-point boundary-value problem (144)-(150) must be solved. During the execution of a sweep, assume that time station t has been reached and that $A_*(t)$, C_*, $\lambda_*(t)$ are known at that time station. Then, Eqs. (145) and

(148) constitute a system of $m + k$ equations in the $m + k$ components of the vectors $B_*(t), \rho_*(t)$. Once more, the system admits a unique solution providing relation (211) is satisfied.

The implication of (211) is that, while the state x and/or the parameter π can be absent from the nondifferential constraint (14), the control u can never be absent. In fact, u must be present in each of the k scalar components of the vector S. Therefore, suitable transformations must be introduced in order to convert problems where the function S does not involve the control into problems where the function S involves the control. For a discussion of these transformations, see Ref. 9.

Starting Condition. The present algorithm can be started with nominal functions $x(t)$ $u(t)$, π satisfying condition (15-1) and violating none, one, or all of conditions (13), (14), (15-2). If the nominal functions are such that Ineq. (189) is satisfied, the algorithm starts with a restoration phase; hence, the first cycle is a half cycle, including a restoration phase only. On the other hand, if the nominal functions are such that Ineq. (190) is satisfied, the algorithm starts with a conjugate gradient phase; hence, the first cycle is a complete cycle, including both a conjugate gradient phase and a restoration phase.

Bypassing Condition. At the end of the conjugate gradient phase of any cycle, the constraint error P must be computed. If Ineq. (189) is satisfied, a restoration phase is started. If Ineq. (190) is satisfied, the restoration phase is bypassed, and the next cycle of the algorithm is started.

Convergence Conditions. For the restoration phase taken individually, convergence is achieved whenever Ineq. (190) is satisfied. For the sequential conjugate gradient-restoration algorithm taken as a whole, convergence is achieved whenever Ineqs. (192) are satisfied simultaneously.

Safeguards. Let N denote the number of iterations. Within each restoration phase, let N_r denote the number of restorative iterations. Within each restorative iteration, let N_{br} denote the number of bisections of the restoration stepsize required to satisfy Ineqs. (194). For the conjugate gradient phase, let N_{bg} denote the number of bisections of the conjugate gradient stepsize required to satisfy Ineqs. (207)-(209). Finally, for a complete conjugate gradient-restoration cycle, let N_{bc} denote the number of bisections of the conjugate gradient stepsize required to satisfy the cycle descent property (210).

With this understanding, the following safeguards are essential to the proper implementation of the sequential conjugate gradient-restoration algorithm:

$$\text{(i)} \qquad N \leq N_* , \tag{212}$$

$$\text{(ii)} \qquad N_r \leq N_{r*} , \tag{213}$$

$$\text{(iii)} \qquad N_{br} \leq N_{br*} , \tag{214}$$

$$\text{(iv)} \qquad N_{bg} \leq N_{bg*} , \tag{215}$$

$$\text{(v)} \qquad N_{bc} \leq N_{bc*} . \tag{216}$$

In the above inequalities, the right-hand sides are specified upper bounds.

Restarting Conditions. The directional coefficient of the conjugate gradient phase must be reset at the level

$$\gamma = 0 \tag{217}$$

if any of these circumstances arise:

$$\text{(i)} \qquad \tilde{J}_\alpha(0) \geq 0 , \qquad \gamma = Q/\hat{Q} , \tag{218}$$

$$\text{(ii)} \qquad N_{bg} \geq 1 , \qquad \gamma = 0 \quad \text{or} \quad \gamma = Q/\hat{Q} , \tag{219}$$

$$\text{(iii)} \qquad N_{bc} \geq 1 , \qquad \gamma = 0 \quad \text{or} \quad \gamma = Q/\hat{Q} , \tag{220}$$

$$\text{(iv)} \qquad N_{bg} > N_{bg*} , \qquad \gamma = Q/\hat{Q} , \tag{221}$$

$$\text{(v)} \qquad N_{bc} > N_{bc*} , \qquad \gamma = Q/\hat{Q} . \tag{222}$$

Satisfaction of Ineq. (218) indicates a loss of the

descent property of the present conjugate gradient phase. Hence, the directional coefficient must be reset at the level (217), characteristic of an ordinary gradient phase.

Satisfaction of Ineq. (219) indicates that the optimum stepsize of the present conjugate gradient phase [this is the stepsize satisfying Ineq. (206)] cannot be employed, owing to violation of one or more of Ineqs. (207)-(209). Hence, this stepsize must be bisected N_{bg} times so as to arrive at satisfaction of (207)-(209), while violating (206). Because of the ensuing large violations of the orthogonality and conjugacy conditions, the directional coefficient of the next conjugate gradient phase must be reset at the level (217), characterisitc of an ordinary gradient phase.

Satisfaction of Ineq. (220) indicates that the optimum stepsize of the present conjugate gradient phase cannot be employed, owing to violation of the cycle descent property (210). Hence, this stepsize must be bisected N_{bc} times, so as to arrive at satisfaction of (210), while violating (206). Once more, because of the ensuing large violations of the orthogonality and conjugacy conditions, the directional coefficient of the next conjugate gradient phase must be reset at the level (217), characteristic of an ordinary gradient phase.

Satisfaction of Ineq. (221) has a stronger implication than satisfaction of Ineq. (219). It requires restarting with $\gamma = 0$ in the present conjugate gradient phase, rather than the next conjugate gradient phase.

Satisfaction of Ineq. (222) has a stronger implication than satisfaction of Ineq. (220). It requires restarting with $\gamma = 0$ in the present conjugate gradient phase, rather than the next conjugate gradient phase.

Nonconvergence Conditions. The sequential conjugate gradient-restoration algorithm must be programmed to stop whenever any of several circumstances arise:

$$\text{(i)} \qquad N > N_* \,, \tag{223}$$

$$\text{(ii)} \qquad N_r > N_{r*} \,, \tag{224}$$

$$\text{(iii)} \qquad N_{br} > N_{br*} \,, \tag{225}$$

$$\text{(iv)} \qquad \tilde{J}_\alpha(0) \geq 0, \qquad \gamma = 0 \,, \tag{226}$$

$$\text{(v)} \qquad N_{bg} > N_{bg*}, \qquad \gamma = 0 \,, \tag{227}$$

$$\text{(vi)} \qquad N_{bc} > N_{bc*}, \qquad \gamma = 0 \,, \tag{228}$$

$$\text{(vii)} \qquad M > M_* \,. \tag{229}$$

Satisfaction of Ineq. (223) indicates extreme slowness of convergence of the algorithm as a whole. Satisfaction of Ineq.

(224) indicates extreme slowness of convergence of the restoration phase. Satisfaction of Ineq. (225) indicates extreme smallness of the restorative displacements. Satisfaction of Ineq. (226) indicates loss of the descent property of the ordinary gradient phase, owing to numerical inaccuracies. Satisfaction of either Ineq. (227) or Ineq. (228) indicates extreme smallness of the ordinary gradient displacements. Finally, satisfaction of Ineq. (229) indicates overflow: the modulus M of some of the quantities used in the algorithm has reached the upper limit M_* allowed by the particular computer employed.

Remark. Several numerical examples illustrating the theory given in this paper are presented in Ref. 16. For a general discussion of the properties of sequential gradient-restoration algorithms, the reader is referred to Ref. 17.

References

1. LASDON, L.S., MITTER, S.K., and WAREN, A.D., The Conjugate Gradient Method for Optimal Control Problems, IEEE Transactions on Automatic Control, Vol. AC-12, No. 2, 1967.

2. HORWITZ, L.B., and SARACHIK, P.E., Davidon's Method in Hilbert Space, SIAM Journal on Applied Mathematics, Vol. 16, No. 4, 1968.

3. LASDON, L.S., Conjugate Direction Methods for Optimal Control, IEEE Transactions on Automatic Control, Vol. AC-15, No. 2, 1970.

4. TRIPATHI, S.S., and NARENDRA, K.S., Optimization Using Conjugate Gradient Methods, IEEE Transactions on Automatic Control, Vol. AC-15, No. 2, 1970.

5. SINNOTT, J.F., Jr., and LUENBERGER, D.G., Solution of Optimal Control Problems by the Method of Conjugate Gradients, Proceedings of the Joint Automatic Control Conference, Philadelphia, Pennsylvania, 1967.

6. HEIDEMAN, J.C., and LEVY, A.V., Sequential Conjugate Gradient-Restoration Algorithm for Optimal Control Problems, Part 1, Theory, Journal of Optimization Theory and Applications, Vol. 15, No. 2, 1975.

7. HEIDEMAN, J.C., and LEVY, A.V., Sequential Conjugate Gradient-Restoration Algorithm for Optimal Control Problems, Part 2, Examples, Journal of Optimization Theory and Applications, Vol. 15, No. 2, 1975.

8. PAGUREK, B., and WOODSIDE, C.M., The Conjugate Gradient Method for Optimal Control Problems with Bounded Control Variables, Automatica, Vol. 4, Nos. 5-6, 1968.

9. MIELE, A., DAMOULAKIS, J.N., CLOUTIER, J.R., and TIETZE, J.L., Sequential Gradient-Restoration Algorithm for Optimal Control Problems with Nondifferential Constraints, Journal of Optimization Theory and Applications, Vol. 13, No. 2, 1974.

10. MIELE, A., PRITCHARD, R.E., and DAMOULAKIS, J.N., Sequential Gradient-Restoration Algorithm for Optimal Control Problems, Journal of Optimization Theory and Applications, Vol. 5, No. 4, 1970.

11. MIELE, A., Method of Particular Solutions for Linear, Two-Point Boundary-Value Problems, Journal of Optimization Theory and Applications, Vol. 2, No. 4, 1968.

12. HEIDEMAN, J.C., Use of the Method of Particular Solutions in Nonlinear, Two-Point Boundary-Value Problems, Journal of Optimization Theory and Applications, Vol. 2, No. 6, 1968.

13. MIELE, A., and IYER, R.R., General Technique for Solving Nonlinear, Two-Point Boundary-Value Problems via the Method of Particular Solutions, Journal of Optimization Theory and Applications, Vol. 5, No. 5, 1970.

14. MIELE, A., and IYER, R.R., Modified Quasilinearization Method for Solving Nonlinear, Two-Point Boundary-Value Problems, Journal of Mathematical Analysis and Applications, Vol. 36, No. 3, 1971.

15. MIELE, A., and CANTRELL, J.W., Study on a Memory Gradient Method for the Minimization of Functions, Journal of Optimization Theory and Applications, Vol. 3, No. 6, 1969.

16. CLOUTIER, J.R., MOHANTY, B.P., and MIELE, A., Sequential Conjugate Gradient-Restoration Algorithm for Optimal Control Problems with Nondifferential Constraints, Part 2, Examples, Rice University, Aero-Astronautics Report No. 127, 1977.

17. MIELE, A., Recent Advances in Gradient Algorithms for Optimal Control Problems, Journal of Optimization Theory and Applications, Vol. 17, Nos. 5/6, 1975.